

A 2F² MEMORY DEVICE SYSTEM AND METHOD

BACKGROUND OF THE INVENTION

The present invention relates to the field of semiconductor manufacture and, more particularly, to a 2F² flash memory. As computers become increasingly complex, the need for improved memory storage increases. At the same time, there is a continuing drive to reduce the size of computers and memory devices. Accordingly, a goal of memory device fabrication is to increase the number of memory cells per unit area.

Memory devices contain blocks or arrays of memory cells. A memory cell stores one bit of information. Bits are commonly represented by the binary digits 0 and 1. A flash memory device is a non-volatile semiconductor memory device in which contents in a single cell or a block of memory cells are electrically programmable and may be read or written in a single operation. Flash memory devices have the characteristics of low power and fast operation making them ideal for portable devices such as laptop or notebook computers, digital audio players and personal digital assistant (PDA) devices.

In flash memory, a charged floating gate is one logic state, typically represented by the binary digit 1, while a non-charged floating gate is the opposite logic state typically represented by the binary digit 0. Charges are injected or written to a floating gate by any number of methods, including avalanche injection, channel injection, Fowler-Nordheim tunneling, and channel hot electron injection, for example. A memory cell or flash memory cell may be characterized in terms of its minimum feature size (F) and cell area (F²). For

example, a standard NOR flash cell is typically quoted as a ten square feature cell and a standard NAND flash cell is approximately a 4.5 square feature cell. Typical DRAM (dynamic random access memory) cells are between 8 F² and 6 F². Cell area (F²) is determined according to a well known methodology and represents the multiple of the number of features along the x and y dimensions of a memory cell. A suitable illustration of feature size is presented in U.S. Patent No. 6,043,562, the disclosure of which is incorporated herein by reference.

Memory devices can be created using 2-dimensional structures or using 3-dimensional structures. The 2-dimensional structures are also referred to as planar structures. Generally, 3-dimensional structures yield smaller cell sizes than planar structures. SRAMs and DRAMs have been designed using 3-dimensional structures, however few flash memory cells are fabricated using 3-dimensional structures. Most flash memory cells are fabricated using planar structures. Some flash memory cells have been fabricated using 3-dimensional structures, but they are, generally, in the size range of 4.5 F² to 8 F² which are not significantly smaller than flash memory cells fabricated using planar structures.

Accordingly, there is a need for a 3-dimensional flash memory device having a cell area of reduced square feature size.

SUMMARY OF THE INVENTION

According to one embodiment of the invention, a memory cell is disclosed. The memory cell comprises a source, a vertical channel, a drain and a horizontal floating gate. The vertical
5 channel is formed over the source. The drain is formed over the vertical channel. The horizontal floating gate is formed over at least a portion of the drain.

According to another embodiment of the invention, a memory cell is disclosed. The memory cell comprises a source, a
10 vertical channel, a drain, a horizontal floating gate and a vertical select gate. The vertical channel is formed over the source. The drain is formed over the vertical channel. The horizontal floating gate is formed over at least a portion of the drain. The vertical select gate is formed perpendicular to the
15 horizontal floating gate.

According to yet another embodiment of the invention, a memory cell is disclosed. The memory cell comprises a first transistor and a select transistor. The first transistor
20 comprises a source, a drain and a gate. The select transistor is coupled to the first transistor and comprises a source, a drain and a gate. The gate of the select transistor is formed perpendicular to the gate of the first transistor.

According to yet another embodiment of the present invention, a memory device is disclosed. The memory device
25 includes a first n-type layer, a p-type layer and a second n-type layer. The p-type layer is formed over the first n-type layer. The second n-type layer is formed over the p-type layer forming a vertical channel.

According to yet another embodiment of the invention, a
30 memory device is disclosed. The memory device includes a

horizontal first n-type layer, a p-type layer, a horizontal second n-type layer, a horizontal floating gate and a vertical select gate. The horizontal first n-type layer is formed over a substrate. The p-type layer is formed over the first n-type layer. The horizontal second n-type layer is formed over the p-type layer. The horizontal floating gate is formed over the substrate. The vertical select gate is formed over the substrate. The p-type layer formed a vertical channel. The first n-type layer forms a buried source and the second n-type layer forms a drain.

According to yet another embodiment of the invention, a memory device is disclosed. The memory device includes a buried source, a vertical channel, a drain, a floating gate and a select gate. The buried source is formed over a substrate. The vertical channel is formed over the buried source. The drain is formed over the vertical channel. The floating gate is formed over the substrate. The select gate is formed perpendicular to the floating gate in a trench formed in the substrate. The memory device has a square feature size of $2F^2$.

According to yet another embodiment of the invention, a memory device is disclosed. The memory device includes a substrate, a first n-type layer, a p-type layer, a second n-type layer, a floating gate, a trench and a select gate. The substrate has at least one semiconductor layer. The first n-type layer is formed over the substrate. The p-type layer is formed over the first n-type layer. The second n-type layer is formed over the p-type layer. The floating gate is formed over the substrate. The trench is formed in the substrate. The select gate is formed on a sidewall of the trench.

- 5 According to yet another embodiment of the invention, a memory device is disclosed. The memory device includes a first n-type layer, a p-type layer, a second n-type layer, a select trench, a vertical select gate, digitlines, a self aligned floating gate and wordlines. The p-type layer is formed over the n-type layer. The select trench is formed in the substrate. The vertical select gate is formed over the second n-type layer. The self aligned floating gate is formed over the n-type layer. The digitlines are formed over the substrate and the digitlines, a select trench, a tunnel oxide layer, a polysilicon layer and an oxide layer. The first n-type layer is formed over the p-type layer. The second n-type layer is formed over the p-type layer. The select gate is formed in the select trench. The tungsten layer is formed over at least a portion of the second n-type layer. The spacer is formed over the tungsten layer, a tunnel oxide layer is formed over the polysilicon layer. The oxide layer is formed on the tunnel oxide layer. The polysilicon layer is formed over the polysilicon layer. A substrate is provided. A first n-type layer is formed over the substrate. A p-type layer is formed over the first n-type layer. A second n-type layer is
- 10
- 15
- 20
- 25
- 30

formed over the p-type layer. A floating gate is formed over the substrate. A trench is formed in the memory device. A select gate is formed in the trench.

According to yet another embodiment of the invention, a
5 method of fabricating a buried source is disclosed. A wafer is provided having a substrate. A periphery of a wafer is covered using an array mask. Source areas are doped with a dopant. An epitaxial deposition is performed to form a p-type channel.

According to another embodiment of the invention, a method
10 of fabricating a memory device is disclosed. A wafer is provided having a substrate. A buried source is formed over the substrate. A vertical channel is formed over the buried source. A cell implant is performed. A tunnel oxide layer is formed over the substrate. A first poly layer is formed over the tunnel
15 oxide layer. A nitride layer is formed over the first poly layer. Wordlines are patterned into the memory device. STI areas are formed in the memory device. The nitride layer is removed. An oxide nitride oxide layer is formed over a surface of the memory device.

According to yet another embodiment of the invention, a
20 method of fabricating a memory device is disclosed. A wafer is provided having a substrate. A buried source is formed over the substrate. A vertical channel is formed over the buried source. A STI area and a self aligned floating gate is formed. A BPSG
25 layer is deposited over the substrate. A hardmask layer is deposited over the BPSG layer. Active areas are patterned to form an active trench. First spacers are formed along sidewalls of the active trench. A drain is formed in the active trench. A wordline is formed over the drain.

According to another embodiment of the invention, a method of fabricating a memory device is disclosed. A buried source is formed in a substrate. A vertical channel is formed over the buried source. A STI area is formed in the memory device. A self aligned floating gate is formed over the substrate. Wordlines are formed over the substrate. A spacer is formed over the wordlines. Rowlines are formed over the substrate. A select gate is formed in a select trench in the substrate.

The methods and devices disclosed, along with variations of them, provide for memory devices having square feature sizes as small as $2F^2$. Such square feature sizes can permit large memory devices, on the order of a gigabyte or larger, to be fabricated on one chip or die. The methods and devices disclosed, along with variations of them, represent a three dimensional fabrication scheme.

BRIEF DESCRIPTION OF THE SEVERAL VIEWS OF THE DRAWINGS

The following detailed description of the present invention can be best understood when read in conjunction with the accompanying drawings, where like structure is indicated with like reference numerals.

Figure 1 illustrates a top view layout of a memory device according to one embodiment of the present invention;

Figure 2A illustrates a cross section of a memory device according to one embodiment of the present invention with reference to line 2A-2A of figure 1;

Figure 2B illustrates a cross section of a memory device according to one embodiment of the present invention with reference to line 2B-2B of figure 1;

Figure 2C illustrates a cross section of a memory device according to one embodiment of the present invention with reference to line 2C-2C of figure 1;

Figure 2D illustrates a cross section of a memory device according to one embodiment of the present invention with reference to line 2D-2D of figure 1;

Figures 3A-3D illustrates a method of fabricating a memory device according to another embodiment of the present invention;

Figure 4 illustrates a top view of a memory device fabricated according to the method of figure 3;

Figure 5 illustrates a portion of a memory device at a selected stage of processing according to the method of figure 3;

Figure 6A illustrates a cross section of a memory device at a selected stage of processing according to the method of figure 3 with reference to line A-A of figure 4;

Figure 6B illustrates a cross section of a memory device at a selected stage of processing according to the method of figure 3 with reference to line B-B of figure 4;

Figure 6C illustrates a cross section of a memory device at a selected stage of processing according to the method of figure 3 with reference to line D-D of figure 4;

Figure 7A illustrates a cross section of a memory device at a selected stage of processing according to the method of figure 3 with reference to line A-A of figure 4;

Figure 7B illustrates a cross section of a memory device at a selected stage of processing according to the method of figure 3 with reference to line B-B of figure 4;

Figure 7C illustrates a cross section of a memory device at a selected stage of processing according to the method of figure 3 with reference to line C-C of figure 4;

Figure 7D illustrates a cross section of a memory device at a selected stage of processing according to the method of figure 3 with reference to line D-D of figure 4;

Figure 8A illustrates a cross section of a memory device at a selected stage of processing according to the method of figure 3 with reference to line A-A of figure 4;

Figure 8B illustrates a cross section of a memory device at a selected stage of processing according to the method of figure 3 with reference to line B-B of figure 4;

Figure 8C illustrates a cross section of a memory device at a selected stage of processing according to the method of figure 3 with reference to line C-C of figure 4;

Figure 8D illustrates a cross section of a memory device at a selected stage of processing according to the method of figure 3 with reference to line D-D of figure 4; and

Attorney Docket No.:
MIO 0065 PA /00-0429

Figure 9 illustrates a computer system in which embodiments of the present invention may be used.

FIG. 9

DETAILED DESCRIPTION OF THE INVENTION

Figure 1 illustrates a top view layout of a memory device 100 according to one embodiment of the present invention. This memory device 100 is generally used for flash memory, but can be used for other types of memory as well. This view illustrates wordlines 104, digitlines 102 and a unit cell or memory cell 101. The unit cell or memory cell 101 is one of many cells of the memory device 100. The memory cell has a minimum feature size of $1F$ or F 105 in a first dimension which is half of the digitline pitch and a feature size of $2F$ 106 in a second dimension which is the wordline pitch. The square feature size or feature area of the cell is thus equal to $2F^2$. The memory cells of this memory device 100 are formed using conventional silicon processing technology. As is described in further detail herein with reference to figures 2A, 2B, 2C and 2D, a select transistor having a select gate 205, source 201 and drain 203 is formed as a part of the memory cell 101. The select gate 205 and a floating gate 206 are formed substantially perpendicular to each other. The select gate 205 of the select transistor and the floating gate 206 make up the minimum feature size of the memory cell 101.

Figure 2A illustrates a cross section of the memory device 100 along the 2A-2A line of figure 1. An n-type layer 201 is formed over a substrate. This n-type layer 201 operates as a source. A p-type layer 202 is formed over the n-type layer 201. The p-type layer 202 can be formed using epitaxial deposition or any other suitable fabrication scheme. One or more drains 203 are formed in the p-type layer 202. A vertical channel 212 is thus created. A select gate 205 is formed for each pair of memory cells of the memory device 100. The select gate 205 is formed vertically.

Digitlines 102 are formed over at least a portion of the drains 203. The digitlines 102 comprise a tungsten layer 210 and a spacer 213 formed over the tungsten layer 210. Additionally, the digitlines 102 may comprise additional layers such as are described in figure 8A. One or more self aligned floating gates 206 are formed horizontally as shown in figure 2A and are perpendicular to the select gates 205. The self aligned floating gates 206 can be fabricated any number of ways such as by forming a first oxide layer over a substrate, a poly layer over the first oxide and a second oxide layer over the poly layer. The self aligned floating gates 206 are sub lithographic features and sub lithographic floating gates. Sub lithographic features are generally created using a removable spacer. Figures 8A, 8B, 8C and 8D illustrate another example of fabricating the self aligned floating gates 206.

Figure 2B illustrates a cross section of the memory device 100 across the 2B-2B line of figure 1. One or more wordlines 104, each comprising a second poly layer 209 and a WSi_x layer 208, are formed over the spacers 213. The spacer 213 is formed of a material selected to insulate the wordlines 104 from the digitlines 102. A shallow trench isolation (STI) area 211 has been formed by etching a trench and depositing a trench oxide layer and filling the trench with oxide. A TiSi layer 221 is formed on the STI area 211 and a TiN layer 220 is formed on the TiSi layer 221 below the tungsten layer 210.

Figure 2C illustrates a cross section of the memory device 100 across the 2C-2C line of figure 1. The vertical select gates 205 are shown. Figure 2D illustrates a cross section of the memory device 100 across the 2D-2D line of figure 1. A

boron-doped phosphosilicate glass (BPSG) layer 214 is formed over the STI area 211. A hardmask 215 is formed over the BPSG 214.

The memory device 100 shown in figures 1, 2A, 2B, 2C and 2D constitutes a $2F^2$ memory cell. It is noted that in fabricating the device 100, removable spacers 216, see figure 2A, may be provided over the floating gates 206 to allow for sublithography to be possible. The removable spacers 216 are merely illustrated with broken lines because they have been removed. Only one removable spacer 216 is shown to preserve clarity. The placement of the select gate reduces over-erasure. Over-erasure is a condition that commonly occurs in flash memory cells in which V_t is caused to go below 0 which causes a transition and conducts or shorts a column of memory cells to ground. Additionally, programming efficiency is increased due to the floating gate 206 being directly above the vertical channel 212.

Figures 3A, 3B, 3C and 3D illustrate a method of fabricating a memory device according to another embodiment of the present invention. An array mask is used to cover a periphery of a wafer at block 301. Buried sources 502, see figure 5, are implanted with a dopant at block 302. The dopant used can be As or Sb. An anneal is performed at block 303. The wafer is cleaned at block 304. The wafer can be cleaned using any number of methods such as by using hydrofluoric acid (HF). An epitaxial deposition (EPI) is performed at block 305 to form a p-type channel 503 of a desired thickness, see figure 5. The desired thickness sets the channel length. The EPI is performed with a dopant such as boron.

Figure 5 illustrates a cross section of the memory device at this stage of processing. Figure 5 shows a p-type substrate 501, buried sources 502 and a p-type channel 503.

Figure 4 is a top level view of a memory device fabricated by the method of figures 3A, 3B, 3C and 3D. The view shows a memory cell 405, wordlines 404 and digitlines 402. The view also shows cross sectional lines A-A, B-B, C-C and D-D which are described in further detail below. Figures 6A-8D illustrate cross sections of a memory device of the present invention at successive points in the fabrication scheme of the present invention.

Referring to figures 6A, 6B, 6C and 3B, a cell implant is performed at block 306. A tunnel oxide layer 604 is formed over a substrate 608 at block 307. A first poly layer 605 is formed over the tunnel oxide layer 604 at block 308. A nitride layer (not shown) is formed or deposited over the first poly layer 605 at block 309. Areas for the wordlines 404 are patterned into the memory device at block 310. The nitride layer, first poly layer 605 and a trench are etched at block 311 to form STI trenches or areas 607. A shallow trench isolation (STI) oxide layer (not shown) is deposited at block 312. The STI oxide layer rounds out the corners of the trench 607. The STI trench 607 is filled with oxide at block 313. The surface of the memory device is polished or planarized using mechanical planarization at block 314. An exemplary type of mechanical planarization which can be used is a chemical mechanical planarization (CMP). The polishing makes the surface of the memory device planar.

The nitride layer is removed at block 315. An oxide nitride oxide (ONO) layer 606 is formed over the surface of the memory device at block 316. Figures 6A, 6B and 6C show the memory device at this stage of the method and, more particularly, show the floating gate 610 and it's alignment to the STI areas 607.

This alignment makes the floating gate 610 a self aligning floating gate.

Figure 6A illustrates a cross section of the memory device in the process of fabrication with reference to the A-A line of figure 4. The tunnel oxide layer 604 is shown formed over the silicon substrate 608. The first poly layer 605 is formed over the tunnel oxide layer 604. The ONO layer 606 is formed over the first poly layer 605. Figure 6B illustrates a cross section of the memory device in the process of fabrication with reference to the B-B and C-C lines of figure 4. This shows how the ONO layer 606 has formed into horizontal and vertical portions. Figure 6C illustrates a cross section of the memory device in the process of fabrication with reference to the D-D line of figure 4 and shows the STI area 607 over the substrate 608.

Referring to figures 3C, 7A, 7B, 7C and 7D, a boron-doped phosphosilicate glass (BPSG) layer 717 is deposited at block 318 over the ONO layer 606. Rapid thermal processing (RTP) is performed on the memory device at block 319. RTP subjects the memory device to a short, controlled thermal cycle. The surface of the memory device is optionally polished by using mechanical planarization again and a hardmask layer 710 is deposited at block 320.

The digitlines or active area 402 of the memory device are patterned at block 321. The digitlines or active area 402 are etched at block 322 down to the tunnel oxide layer 604 to form a trench or active trench 718. The hardmask layer 710, BPSG layer 717, ONO layer 606 and first poly layer 605 of the trench 718 are etched away, but the tunnel oxide layer 604 is not etched. A first spacer layer is deposited and etched at block 323 to vertically form first spacers 711. Drains 714 are formed in the

active areas or columns by implanting a dopant at block 324. Another RTP is performed at block 325. TiSi 713 and TiN 712 layers are formed over the drains 714 at block 326. The TiN 712 and TiSi 713 layers are formed horizontally and vertically in the active trench 718. Another RTP is performed at block 327. A tungsten layer 716 is deposited over the active areas or columns in the active trench 718 at block 328. Mechanical planarization is performed on the memory device so that the tungsten layer 716 is planar with the hardmask at block 329. The tungsten layer 716 is etched such that approximately half is removed at block 330. Second spacers 715 are deposited over the tungsten layer 716 at block 331. The second spacers 715 fill the rest of the trench so the height of the active area or columns is approximately equal to the height of the hard mask 710.

The digitlines 402 comprise the second spacers 715 and the tungsten layer 716. The digitlines 402 are insulated because of the second spacers 715. Figures 7A, 7B, 7C and 7D illustrate the formation of digitlines 402. Figure 7A shows a cross section of the memory device in the process of fabrication with reference to the A-A line of figure 4. The BPSG layer 717 is formed over the ONO layer 606. The hardmask 710 is formed over the BPSG layer 717. The first spacers 711 are formed vertically adjacent to the BPSG layers after the trench 718 has been etched away. The tungsten layer 716 is formed over the Ti layers, TiN 712 and TiSi 713. The second spacers 715 are formed over the tungsten layer 716 in the trench or active areas 718. Figure 7B shows a cross section of the memory device in the process of fabrication with reference to the B-B line of figure 4. Figure 7C shows a cross section of the memory device in the process of fabrication with reference to the C-C line of figure 4. Figure 7D shows a cross

section of the memory device in the process of fabrication with reference to the D-D line of figure 4.

Referring to figures 3D, 8A, 8B, 8C and 8D, the hardmask layer 710 and BPSG layer 717 are removed or etched from the wordlines 404 at block 332. A removable spacer 825 is deposited at block 333. Only one removable spacer is shown in the figures to preserve clarity. The removable spacer 825 is etched at block 334. At least one select trench 820 is formed by etching the ONO layer 606, first poly layer 605, the tunnel oxide 604 and silicon to a desired depth at block 335. The remaining portion of the removable spacer 825 is removed at block 336. A select transistor oxide layer 822 is formed on the surface of the select trench 820. A second poly layer 821 is formed over the surface of the memory device, including the select trench 820 and a WSi_x layer 823 is deposited over the second poly layer 821 at block 338. The second poly layer 821 is also referred to as the wordline poly. The second poly layer 821 and WSi_x layer 823 are patterned at block 339 and etched at block 340. By etching and removing the removable spacer 825, the second poly layer 821 and floating gate 605 are capacitively coupled. Figures 8A, 8B, 8C and 8D show wordline 404 formation. Figure 8A is a cross section of the memory device in the process of fabrication with reference to the A-A line of figure 4. The select trenches 820 have a layer of select gate oxide 822 and are filled with the second poly layer 821. The removable spacer 825 has been removed. The second poly layer 821 is shown in the select trenches 820 and other areas. Figure 8B is a cross section of the memory device in the process of fabrication with reference to the B-B line of figure 4. The wordlines 404 are shown and comprise the WSi_x layer 823 formed over the second poly layer 821 formed over the

second spacer 715. Thus, the rowlines 404 are insulated from the tungsten layer 716 by the second spacer 715. Figure 8C illustrates a cross section of the memory device in the process of fabrication with reference to the C-C line of figure 4. The select trenches 820 are shown. Figure 8D illustrates a cross section of the memory device in the process of fabrication with reference to the D-D line of figure 4.

Figure 9 is an illustration of a computer system 912 that can use and be used with embodiments of the present invention. The computer system can be a desktop, network server, handheld computer or the like. As will be appreciated by those skilled in the art, the computer system 912 would include ROM 914, mass memory 916, peripheral devices 918, and I/O devices 920 in communication with a microprocessor 922 via a data bus 924 or another suitable data communication path. The memory devices 914 and 916 can be fabricated according to the various embodiments of the present invention, including memory devices having a square feature size of $2F^2$. ROM 914 can include EPROM or EEPROM or flash memory. Mass memory 916 can include DRAM, synchronous RAM or flash memory.

The present inventor recognizes that other 3-dimensional memory cells place the floating gate in the sidewall of a trench in the $\langle 111 \rangle$ plane or other planes which have a higher density of bonds. This placement typically results in an inferior oxide resulting in retention, cycling and trapping problems with the memory cell. The present invention generally places the floating gate in the $\langle 100 \rangle$ plane thereby avoiding the aforementioned results.

For the purposes of describing and defining the present invention, formation of a material "on" a substrate or layer

refers to formation in contact with a surface of the substrate or layer. Formation "over" a substrate or layer refers to formation above or in contact with a surface of the substrate. A "flash memory device" includes a plurality of memory cells. Each
5 "memory cell" of a flash memory device can comprise components such as a gate, floating gate, control gate, wordline, channel region, a source, self aligned source and a drain. The term "patterning" refers to one or more steps that result in the removal of selected portions of layers. The patterning process
10 is also known by the names photomasking, masking, photolithography and microlithography. The term "self-aligned gate" refers to a memory device where the gate electrodes are formed before the source/drain diffusions are made. An "anneal" is a high temperature processing step designed to minimize stress in the crystal structure of the wafer. An "epitaxial deposition"
15 (EPI) involves depositing a layer of high-quality, single-crystal silicon on a wafer surface to form a base. The term "rapid thermal processing (RTP)" refers to a process that subjects a wafer to a short, yet controlled, thermal cycle which heats the
20 wafer from room temperature to a high temperature, such as 1200°C, in a few seconds.

Many other electronic devices can be fabricated utilizing various embodiments of the present invention. For example, memory devices according to embodiments of the invention can be
25 used in electronic devices such as cell phones, digital cameras, digital video cameras, digital audio players, cable television set top boxes, digital satellite receivers, personal digital assistants and the like. Additionally, large capacity flash memory chips can be fabricated. For example, a $0.45\mu^2$ cell can be
30 realized in 0.15μ technology using a $2F^2$ memory cell.

Having described the invention in detail and by reference to preferred embodiments thereof, it will be apparent that modifications and variations are possible without departing from the scope of the present invention defined in the appended
5 claims. Other suitable materials may be substituted for those specifically recited herein. For example, the substrate may be composed of semiconductors such as gallium arsenide or germanium. Additionally, other dopants may be utilized besides those specifically stated. Generally, dopants are found in groups III
10 and V of the periodic table.